JOURNAL OF GEOPHYSICAL RESEARCH: ATMOSPHERES, VOL. 118, 355–369, doi:10.1029/2012JD018590, 2013

Evaluating the impact of orbital sampling on satellite–climate model comparisons

Bin Guan,[1,2] Duane E. Waliser,[1] Jui-Lin F. Li,[1] and Arlindo da Silva[3]

Received 30 July 2012; revised 14 November 2012; accepted 15 November 2012; published 24 January 2013.

[1] The effect of orbital sampling is one of the chief uncertainties in satellite–climate model comparisons. In the context of an ongoing activity to make satellite data more accessible for model evaluation (i.e., obs4MIPs), six variables (temperature, specific humidity, ozone, cloud water, cloud cover, and ocean surface wind) associated with six satellite instruments are evaluated for the orbital sampling effect. Comparisons are made between reanalysis and simulated satellite-sampled data in terms of bias and pattern similarity. It is found that the bias introduced by orbital sampling for long-term annual means, monthly climatologies, and monthly means is largely negligible, which is within ~3% of the standard deviation of the three quantities for most fields. The bias for 2-hPa temperature and specific humidity, while relatively large (9–10%), is within the estimated observational uncertainty. In terms of pattern similarity, cloud water and upper level specific humidity are the most sensitive to orbital sampling among the variables considered, with the magnitude of the sampling effect dependent on the spatial resolution—insignificant at $1.25^\circ \times 1.25^\circ$ resolution for both. For all variables considered, orbital sampling effects are not an important consideration for model evaluation at $1.25^\circ \times 1.25^\circ$ resolution. At $0.5^\circ \times 0.5^\circ$, orbital sampling is potentially important for cloud water and upper level specific humidity when evaluating model long-term annual means and monthly climatologies, and for cloud water when evaluating monthly means, all in terms of pattern similarities. Orbital sampling is not an important factor for evaluating zonal means in call cases considered.

Citation: Guan, B., D. E. Waliser, J. F. Li, and A. da Silva (2013), Evaluating the impact of orbital sampling on satellite–climate model comparisons, *J. Geophys. Res. Atmos.*, *118*, 355–369, doi:10.1029/2012JD018590.

1. Introduction

[2] The wealth of information provided by Earth-observing satellites has greatly advanced our understanding of the earth system and our ability to forecast its behavior. An important use of satellite observations is the evaluation of global climate models. Compared to observations from other platforms, satellite observations have the advantage of increased spatiotemporal coverage, which facilitates model evaluation on a global scale. Currently, active satellite missions are monitoring most components of the climate system. For some variables, nearly a decade of satellite observations have been accumulated (e.g., the A-Train observations). For many with an operational component, the length of satellite record is up to two decades or more (e.g., outgoing longwave radiation [OLR], sea level, sea surface temperature [SST], etc.). These will provide more and improved opportunities for satellite-based model evaluation in the coming decades.

[3] A practical difficulty in the evaluation of global models with satellite data concerns the particular space-time sampling of the satellites. For example, a polar-orbiting satellite samples the globe with gaps in time and space determined by the orbital character, instrument swath, etc., and a geostationary satellite provides relatively continuous monitoring in time but only for a certain portion of the globe. For evaluation of climate models, satellite observations will be most useful if provided in a manner that is compatible with the model output. Under the recently initiated obs4MIPs project [*Gleckler et al.*, 2011; *Teixeira et al.*, 2011; http://obs4mips.llnl.gov:8080/wiki/] which aims to make satellite-based (and other observationally-based) data more accessible for climate model evaluation, the National Aeronautics and Space Administration (NASA), in collaboration with the Department of Engergy (DOE) and the National Oceanic and Atmospheric Administration (NOAA), is preparing a suite of well established satellite products that closely follow the Coupled Model Intercomparison Project Phase 5 (CMIP5) simulation protocol [*Taylor et al.*, 2012]. Eighteen such products have been made available to the science community through the Earth System

[1]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA.
[2]Joint Institute for Regional Earth System Science and Engineering, University of California, Los Angeles, California, USA.
[3]NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.

Corresponding author: B. Guan, Jet Propulsion Laboratory, MS 233-300, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA. (bin.guan@jpl.nasa.gov)

2169-897X/13/2012JD018590

Grid Federation [ESGF; *Williams et al*., 2009], which officially hosts the CMIP5 model output.

[4] Efforts like obs4MIPs make satellite–model comparisons easier. On the other hand, it remains to be fully understood to what extent the particular space-time sampling of the atmosphere by satellites as opposed to the regular and continuous sampling by global models affects the results and interpretation of the comparison between satellite observations and model simulations. Conceptually, the effect of orbital sampling can be evaluated by comparing the "true" geophysical state and the satellite-sampled state. When lacking ground-truth observations, the former has often been approximated either by statistical models [*Laughlin*, 1981; *North et al*., 1993; *Bell and Kundu*, 1996; *Soman et al*., 1996] or general circulation models [GCMs; *Engelen et al*., 2000; *Lin et al*., 2002; *Luo et al*., 2002; *Li et al*., 2007; *Aghedo et al*., 2011]. These studies documented the orbital sampling effect for selected satellites/instruments (TIROS/TOVS, TRMM/TMI, TRMM/PR, TRMM/VIRS, TRMM/CERES, Aura/TES, Aura/MLS) and geophysical variables (brightness temperature, precipitation, OLR, ozone, carbon monoxide, cloud ice, water vapor, temperature). These studies were often based on observations from a limited area and/or a short period of time. Since the difference between the complete spatiotemporal field and the satellite-sampled field is a function of not only the orbital geometry but also the statistical characteristics of the geophysical variable itself [*Bell et al*., 1990; *Li et al*., 1996], it is not clear to what extent the results based on a limited area can be applied to other areas of the world [*North*, 1988], and to what extent the evaluation based on modeled atmosphere represents the sampling effect on real atmosphere. As such, more comprehensive assessments are needed, especially in the context of ongoing international efforts for model evaluation (e.g., CMIP5) and making satellite observations more suitable for doing that (e.g., obs4MIPs). In this regard, a question of key interest that remains to be addressed is whether orbital sampling affects the statistical characteristics of the long-term annual mean, monthly climatology (i.e., the 12-month annual cycle), and monthly means (over a relatively long period of time) of a geophysical variable. Both bias and pattern similarity between the complete and satellite-sampled data are of interest since these two aspects are both important and routinely examined when evaluating climate models [e.g., *Randall et al*., 2007]. It is noted that most previous studies were focused on the bias or total root-mean-square (RMS) error of the sampled data, leaving quantitative comparison of pattern similarity largely unexplored.

[5] Reanalysis data are used in the current study as a more realistic approximation of the real atmosphere for evaluating the orbital sampling effect. In the context of CMIP5 and obs4MIPs, six variables associated with six NASA satellite instruments are considered in our analysis, which represent a cross section of variables that are well established and documented, and useful to the modeling and model evaluation community. Consideration of a variety of satellite instruments and variables in the same analysis also facilitates direct comparisons of their relative sensitivity to orbital sampling, an aspect lacking in previous studies. A variable of particular interest is cloud water, which critically affects radiative budgets of the atmosphere [e.g., *Waliser et al*., 2011], yet bears a key shortcoming in global climate models [e.g., *Waliser et al*., 2009; *Li et al*., 2012a, 2012b]. The evaluation of modeled cloud water has become possible relatively recently using satellite observations [*Li et al*., 2005, 2007, 2012a, 2012b; *Su et al*., 2006; *Jiang et al*., 2010, 2012], for which the effect of orbital sampling remains to be explored, especially relative to other variables of interest.

[6] Several obs4MIPs products involve merging of data from different observational platforms. For example, the CERES top-of-atmosphere radiative fluxes for CMIP5 merges observations from sun-synchronous satellite and multiple geostationary satellites for improved representation of the diurnal cycle (http://ceres.larc.nasa.gov/cmip5_data.php). Similarly, TRMM 3B42 precipitation is a merged product based on many sources for improved spatial coverage and data quality [*Huffman et al*., 2007]. The variables related to these data sets are not considered in the current analysis since the comparison of these data sets to models not only depends on orbital characters of the satellites, but also the particular technique used to merge data from various platforms.

2. Data and Methods

2.1. Data

[7] Reanalysis data with sub-daily time steps are used as an approximation of the real atmosphere. Noting that each reanalysis product has its own biases and limitations, two reanalysis products, namely, the National Centers for Environmental Prediction (NCEP) Climate Forecast System Reanalysis [CFSR; *Saha*, 2010] and NASA Modern-Era Retrospective Analysis for Research and Applications [MERRA; *Rienecker et al*., 2011], are separately analyzed, with the difference between the two products serving as a rough measure of the current observational uncertainty. CFSR is an operational, real-time product that features the coupling between the atmosphere, ocean, land surface, and sea ice components during data assimilation. MERRA is a near-real-time product with a special aim to improve the representation of the global hydrological cycle. While many aspects of the two reanalysis products are similar, differences are substantial for precipitation and other cloud-related variables that are still poorly constrained in these systems [*Rienecker et al*., 2011]. Six variables from each reanalysis product are considered, which are vertical profiles of temperature, specific humidity, ozone mixing ratio, cloud water (ice + liquid) mixing ratio, column total cloud cover, and ocean surface wind. The CFSR product used has a 6-hour time step, with a horizontal resolution of $0.5^\circ \times 0.5^\circ$. The MERRA product used has a 3-hour time step, with a $1.25^\circ \times 1.25^\circ$ resolution. For reference, the CMIP5 coupled models have a spatial resolution of ~0.5–4° for the atmospheric component [*Taylor et al*., 2012]. Nine standard pressure levels between 1000- and 2-hPa inclusive are extracted for all vertical profiles except CFSR cloud water for which one less level (i.e., 2-hPa) is available. The main analysis covers the period of 2000–2010. Sensitivity to analysis period and spatiotemporal resolution will be addressed.

2.2. Sampling Method

[8] Each of the above variables is associated with one or more NASA satellite instruments that enter the ESGF (http://esg-datanode.jpl.nasa.gov) under obs4MIPs. The

Table 1. Variables and Corresponding Satellites/Instruments to be Assessed for Orbital Sampling Effect

Variable	Satellite	Instrument (Swath)
Temperature & Specific Humidity: ≥300 hPa	Aqua	AIRS (1650 km)
Temperature & Specific Humidity: <300 hPa	Aura	MLS (7 km)
Ozone Mixing Ratio	Aura	TES (5.3 km)
Cloud Water Mixing Ratio	CloudSat	CPR (1.4 km)
Column Total Cloud Cover	Aqua	MODIS (2330 km)
Ocean Surface Wind	QuikSCAT	SeaWinds (1800 km)

orbital sampling effect is evaluated for each variable for one or two satellite instruments that are considered the most relevant for the given variable, as listed in Table 1. Two instruments listed (AIRS, MODIS) are aboard the Aqua satellite, two (MLS, TES) are aboard Aura, one (CPR) on CloudSat, and one (SeaWinds) is carried by QuikSCAT. Aqua, Aura, and CloudSat fly a few minutes apart in sun-synchronous orbits, cross the equator around 1:30 pm and 1:30 am local standard time, and complete about 15 orbits per day. They form the A-train satellite constellation [*L'Ecuyer and Jiang*, 2010] together with other satellites. QuikSCAT also flys in sun-synchronous orbits, with equator crossing time around 6:00 pm and 6:00 am, and completes about 14 orbits per day. Figure 1a shows the A-Train and QuikSCAT orbits over a given 24-hr period.

[9] Satellite sampling is simulated for the above instruments with the reanalysis data as input. The actual satellite products are not used. For each reanalysis time step t at interval Δt, orbital sampling is simulated by setting the grid cells not viewed by a particular instrument during $t-\Delta t/2$ and $t+\Delta t/2$ to undefined, considering both the location of the satellite and the swath width of a particular instrument. For a grid cell to be viewable, at least part of its area should fall within the swath of an instrument. Monthly means are then obtained from the complete reanalysis data (hereafter, raw data) and simulated satellite-sampled data, respectively, from which the monthly climatology (i.e., the long-term mean of each month) and long-term annual mean is further obtained.

[10] For a given month and a given pixel, the difference between raw and sampled data will be dependent on the number of reanalysis time steps sampled over the given month [e.g., *Lin et al.*, 2002]. Figure 1b shows the zonally-averaged frequency (percent 6- or 3-hourly time steps) each reanalysis pixel is sampled over a given month. Sampling frequencies are considerably high for instruments with big swaths (Figure 1b, lower 3 panels), and in general higher toward the polar regions. For instruments with single footprint swaths (Figure 1b, upper 3 panels), sampling frequencies are notably low (<2% in the tropics and extratropics). Figure 1b suggests that for instruments with swaths much smaller than the grid size, sampling frequency is dependent on both the time interval and grid size of the data being sampled (Figure 1b, upper 3 panels). In this regard, the coarser spatial resolution of MERRA (by a factor of 2.5) more than compensates for the higher temporal resolution (by a factor of 2), resulting in a 25% higher sampling frequency than for CFSR. For instruments with swaths much larger than the grid size (Figure 1b, lower 3 panels), sampling frequency becomes insensitive to the grid size; in this case, sampling frequency for CFSR is twice that for MERRA due to the factor of 2 difference in the temporal resolution. These factors need to be taken into account when comparing the sampling effect between CFSR and MERRA.

[11] Examples of raw and satellite-sampled data are shown in Figure 2 (left 2 columns) for an arbitrary month and selected CFSR fields. Raw and sampled fields are very similar for AIRS temperature and specific humidity profiles, MODIS cloud cover, and QuikSCAT ocean surface wind. For MLS and TES, satellite-sampled data capture the broad

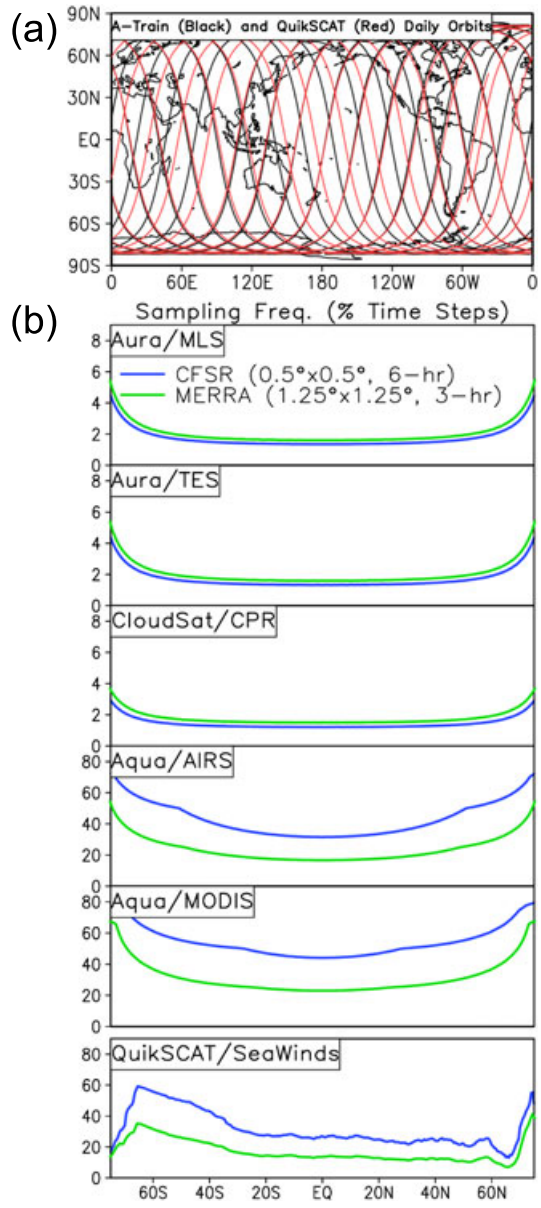

Figure 1. (a) A-Train and QuikSCAT orbits over a given 24-hr period. (b) Frequency (percent 6- or 3-hourly time steps) at which each CFSR (blue) and MERRA (green) grid cell is sampled by various satellite instrument during a given 31-day period (i.e., 124 time steps for CFSR, and 248 time steps for MERRA), shown in zonal averages as a function of latitude. For QuikSCAT, sampling frequency is counted over ocean grids only. Note larger scales for the lower three panels.

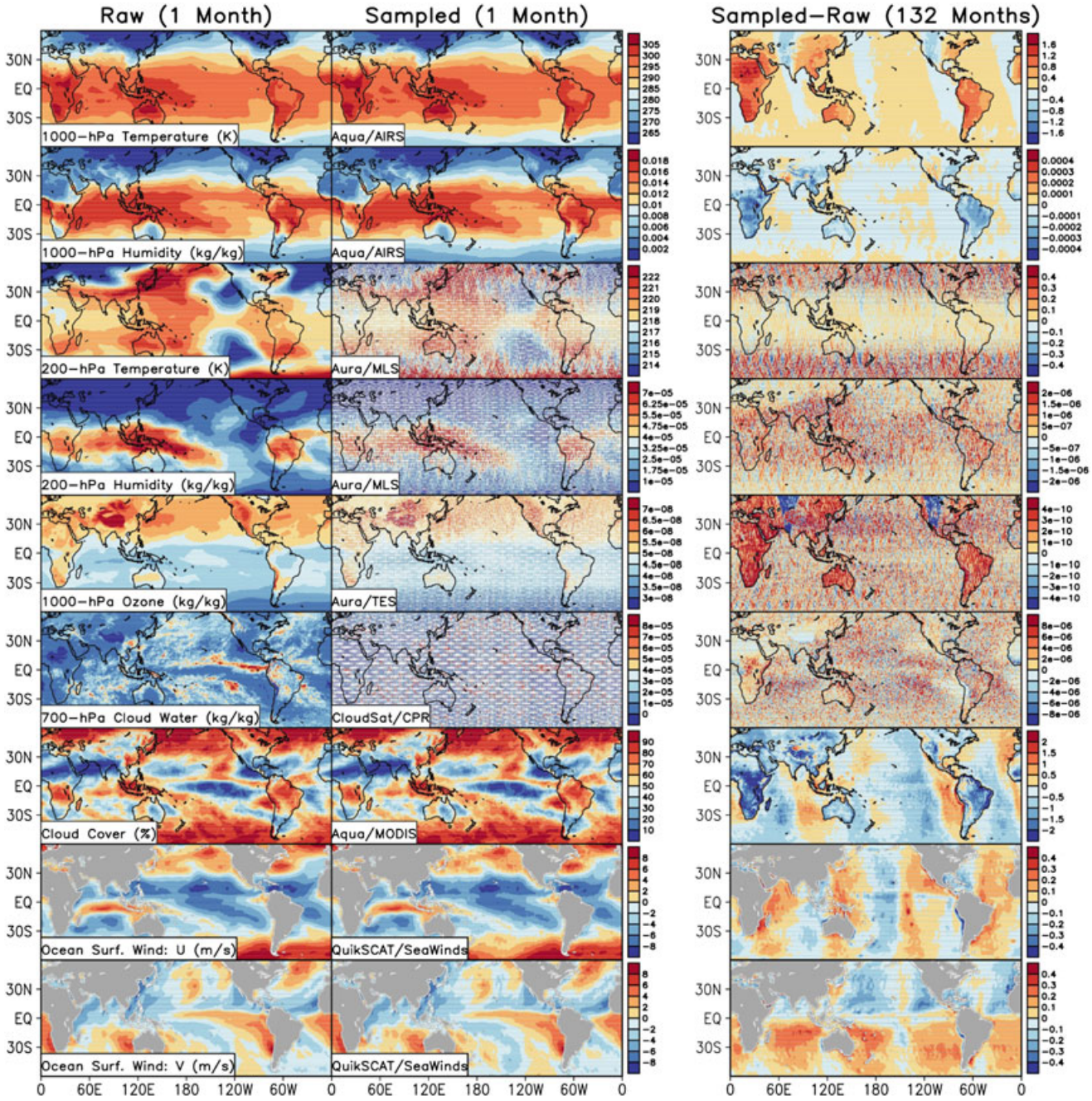

Figure 2. (left 2 columns) Example monthly means over an arbitrary month based on (left) raw and (center) satellite-sampled 6-hourly CFSR fields. (right) Mean difference between sampled and raw data over the period of 2000 –2010. Each row is for a different variable. For vertical profiles, only one representative level is shown for each satellite instrument. Cloud water includes both ice and liquid water throughout this paper.

scale patten of the fields, but not the details. The visually largest difference is for CloudSat/CPR. For each of these instruments, the degree of closeness between raw and sampled data is consistent with the swath width (Table 1) and hence sampling frequencies (Figure 1b).

2.3. Evaluation Method

[12] The effect of orbital sampling is evaluated for the long-term annual mean, monthly climatology, and monthly means, respectively, based on raw and satellite-sampled data. Raw and sampled data are compared in terms of bias (i.e., overall difference over time and space), and pattern similarity represented by standard deviation ratio (sampled data standard deviation divided by raw data standard deviation), correlation, and centered RMS error (i.e., RMS error between two fields after they are centered by their respective means). Each two of the latter three quantities determine the other one, and all three can be shown succinctly in Taylor

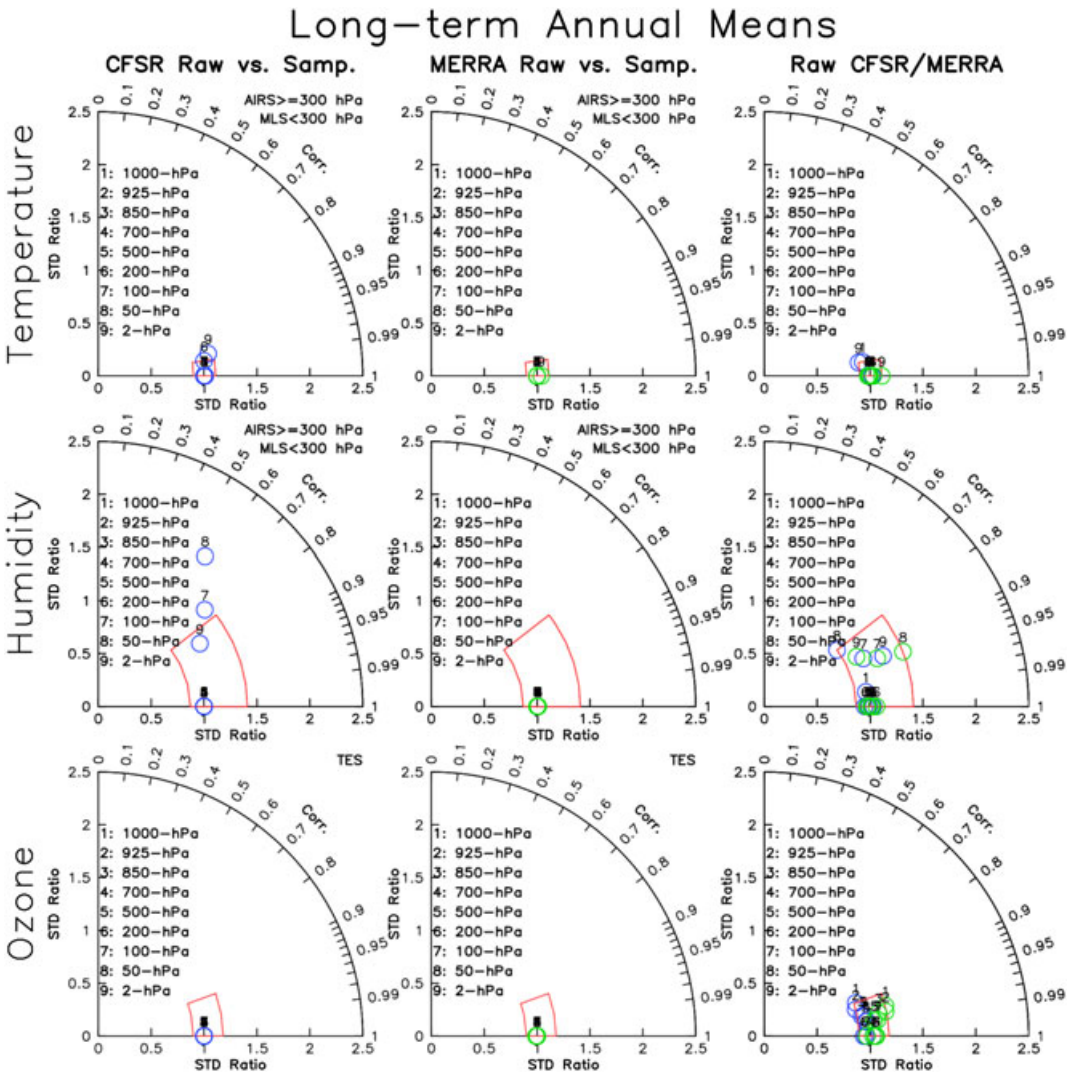

Figure 3. Taylor diagrams showing the comparison of the long-term annual means between raw (the reference point) and satellite-sampled fields based on (left) CFSR and (center) MERRA, and (right) between raw CFSR/MERRA and their mean (the reference point). Each row is for a different variable. The red "box" in each panel represents the uncertainty of the reanalysis data determined by standard deviation ratio and correlation ranges shown in the right column. For ocean surface wind, the statistics are first calculated separately for the zonal and meridional components, and then averaged. Note different scale for cloud water.

diagrams [*Taylor*, 2001]. In the Taylor diagrams to be shown (e.g., Figure 3), the distance between any point in the plot and the origin point indicates the standard deviation ratio of sampled to raw data (abscissa and ordinate), the azimuthal angle of any point relates to the correlation between raw and sampled data (scale on the arc), the point with standard deviation ratio one and correlation one (the reference point) corresponds to the raw reanalysis data, and the distance between any point and the reference point indicates the centered RMS error. As can be shown, centered RMS error and bias add quadratically to the total RMS error [*Taylor*, 2001].

3. Results

3.1. Long-Term Annual Means

[13] Long-term annual means are compared between raw and sampled data for six variables in Figure 3 in terms of pattern similarity. Results based on CFSR and MERRA are shown in the left and center columns, respectively. Since the reanalysis data themselves are subject to uncertainty, which needs to be quantified, the right column compares raw CFSR/MERRA to their mean. The red "box" in each panel delineates the range of the standard deviation ratios and correlations shown in the right column, as a rough estimate of the observational uncertainty [e.g., *Li et al*., 2012a]. In this regard, the difference between raw and sampled data will be deemed as significant if larger than the estimated observational uncertainty.

[14] Based on the above criteria, temperature and ozone profiles, column total cloud cover, and ocean surface wind show insignificant difference between raw and sampled data. For these variables, the uncertainty of the reanalysis data themselves are also relatively small. Significant differences between raw and sampled data are seen for CFSR specific humidity at 100- and 50-hPa levels and for CFSR cloud water at the 100-hPa level, which are contributed by both

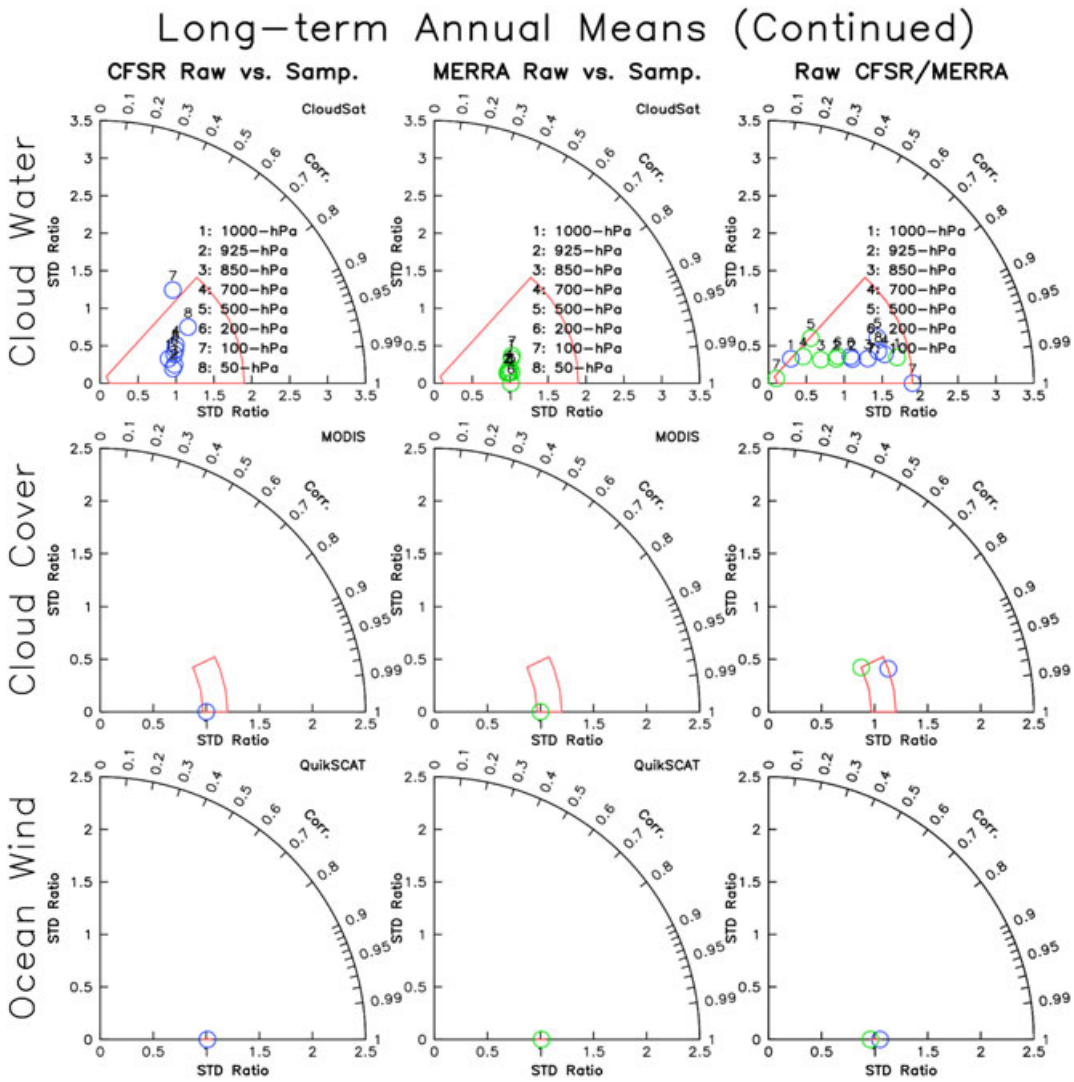

Figure 3. (continued)

differing magnitudes of standard deviations and weak correlations. For these two variables, the difference between raw CFSR and MERRA is also considerably larger than for other variables. The sampling effect is consistently larger with CFSR, which is most likely due to the finer spatial resolution (to be addressed in section 4).

[15] Zonal means are often calculated when evaluating climate models [e.g., *Randall et al*., 2007]. Zonal mean specific humidity is used here to illustrate the orbital sampling effect in this regard. Specific humidity is considered due to its relatively large sensitivity to orbital sampling as seen in Figure 3. The observational uncertainty (i.e., the red "box") remains nearly unchanged for the zonal means (cf. Figure 4 and 2nd row of Figure 3). On the other hand, the sampling effect becomes much smaller and insignificant (Figure 4).

3.2. Monthly Climatologies

[16] Similar comparisons for monthly climatologies (i.e., long-term means of each month) are shown in Figure 5. Much as in the case of long-term annual means, temperature and ozone profiles, column total cloud cover, and ocean surface wind show insignificant difference between raw and sampled data. For these variables, the uncertainty of the reanalysis data themselves are also relatively small. Significant differences between raw and sampled data are seen for CFSR specific humidity at 100- and 50-hPa levels and for CFSR cloud water at 700-, 500-, 200-, and 100-hPa levels. For these two variables, raw CFSR and MERRA also differ significantly. The sampling effect is, again, consistently larger with CFSR.

[17] Next, zonal mean cloud water is examined. Cloud water is considered due to its relatively large sensitivity to orbital sampling as seen in Figure 5. The overall sampling effect is insignificant compared to the observational uncertainty (Figure 6a). A more detailed examination is given by the latitude-height cross sections (Figure 6b). The orbital sampling effect in this case is dominated by different standard deviations between the raw and sampled data (Figure 6b, left panels), as correlations are in general high (Figure 6b, center panels) and biases low (Figure 6b, right panels). Beside height-dependence, as already seen in Figure 5, latitudinal dependence is evident in Figure 6b (left panels). Both CFSR and MERRA show largest sensitivity to orbital sampling at the

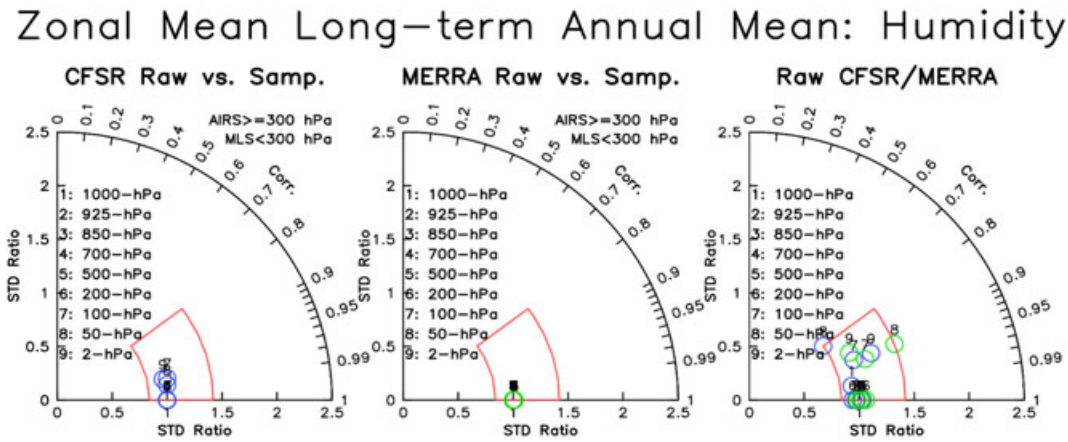

Figure 4. As the 2nd row in Figure 3 except for the zonal mean field.

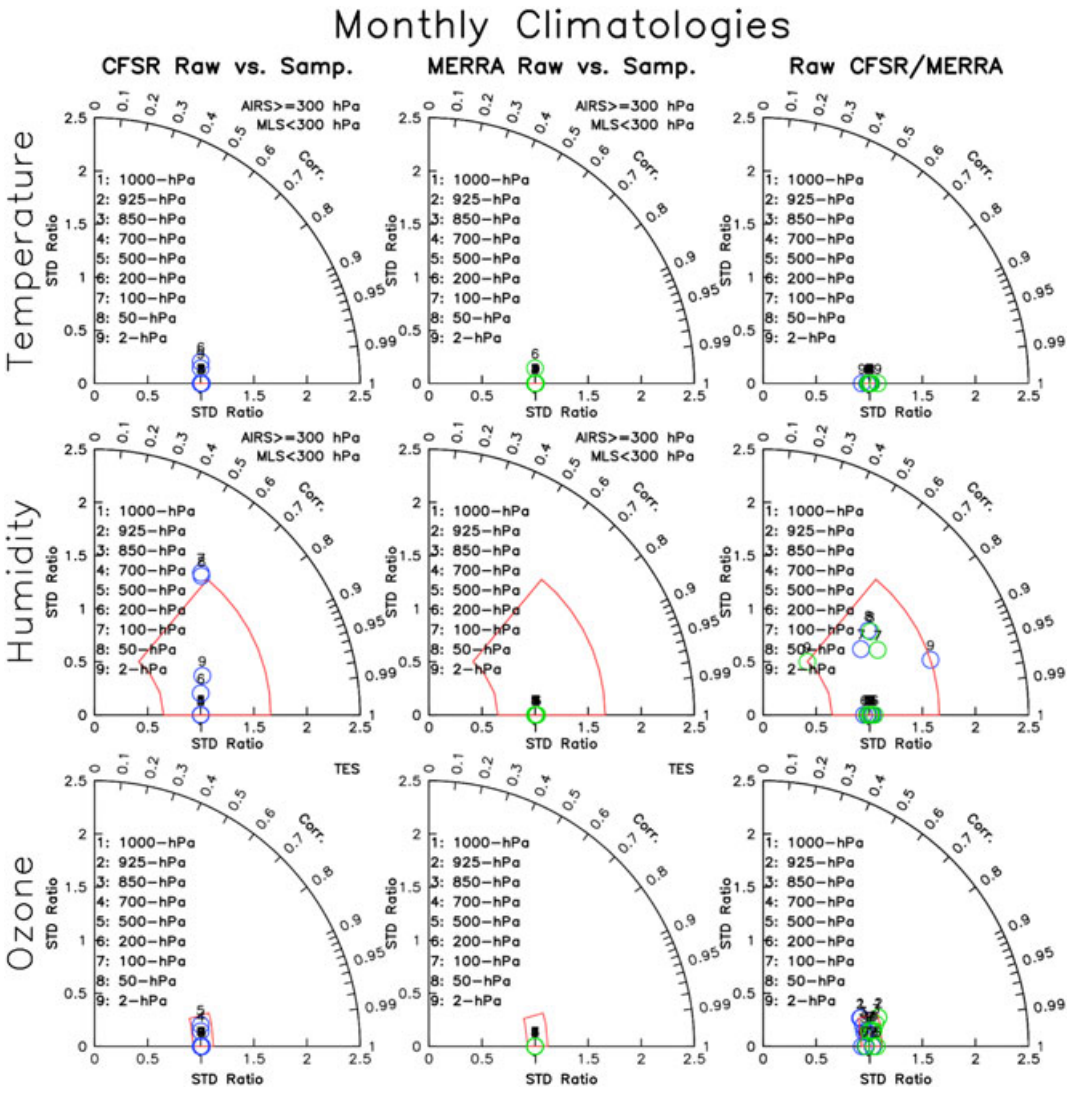

Figure 5. As Figure 3 except for monthly climatologies.

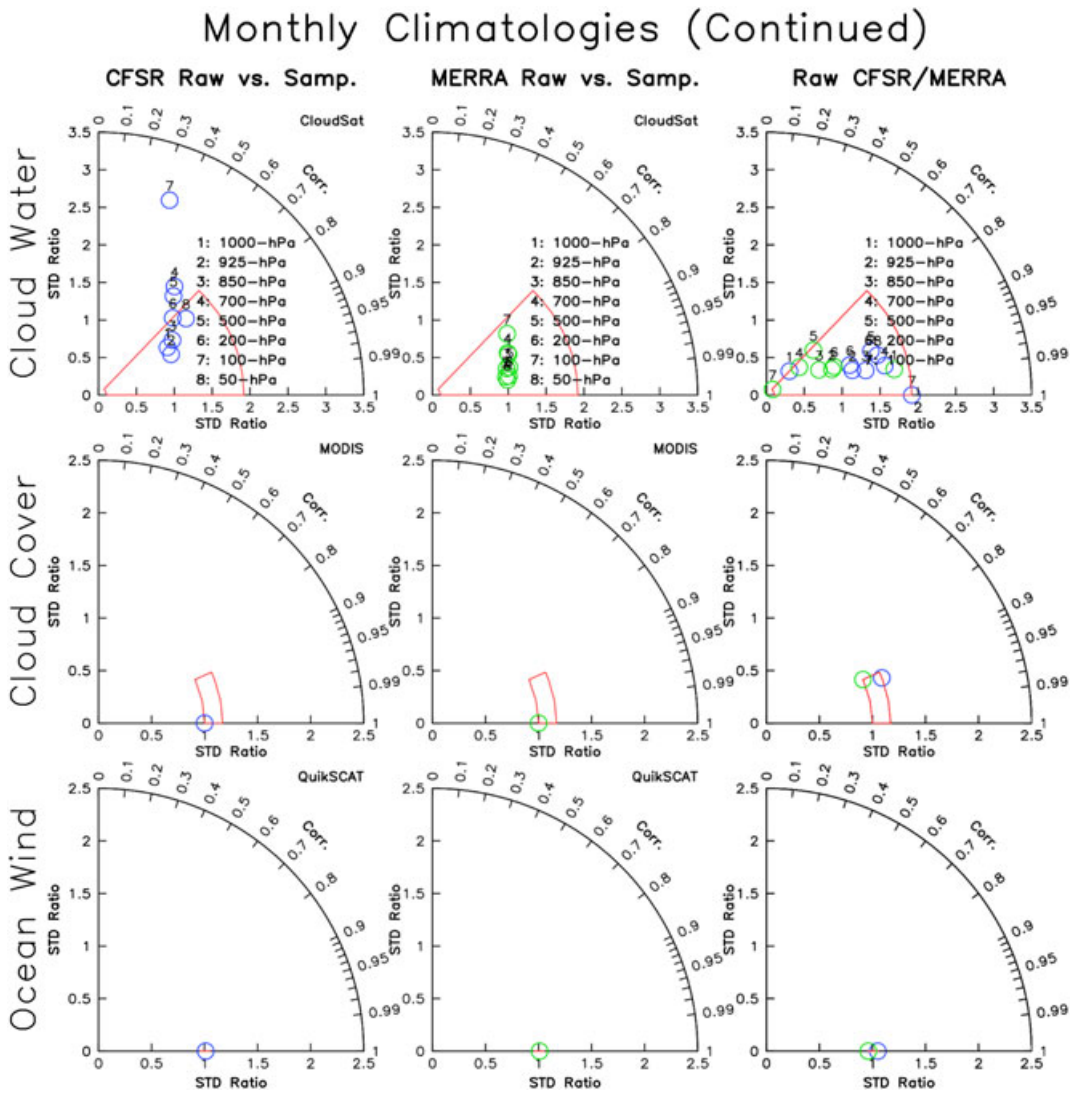

Figure 5. (continued)

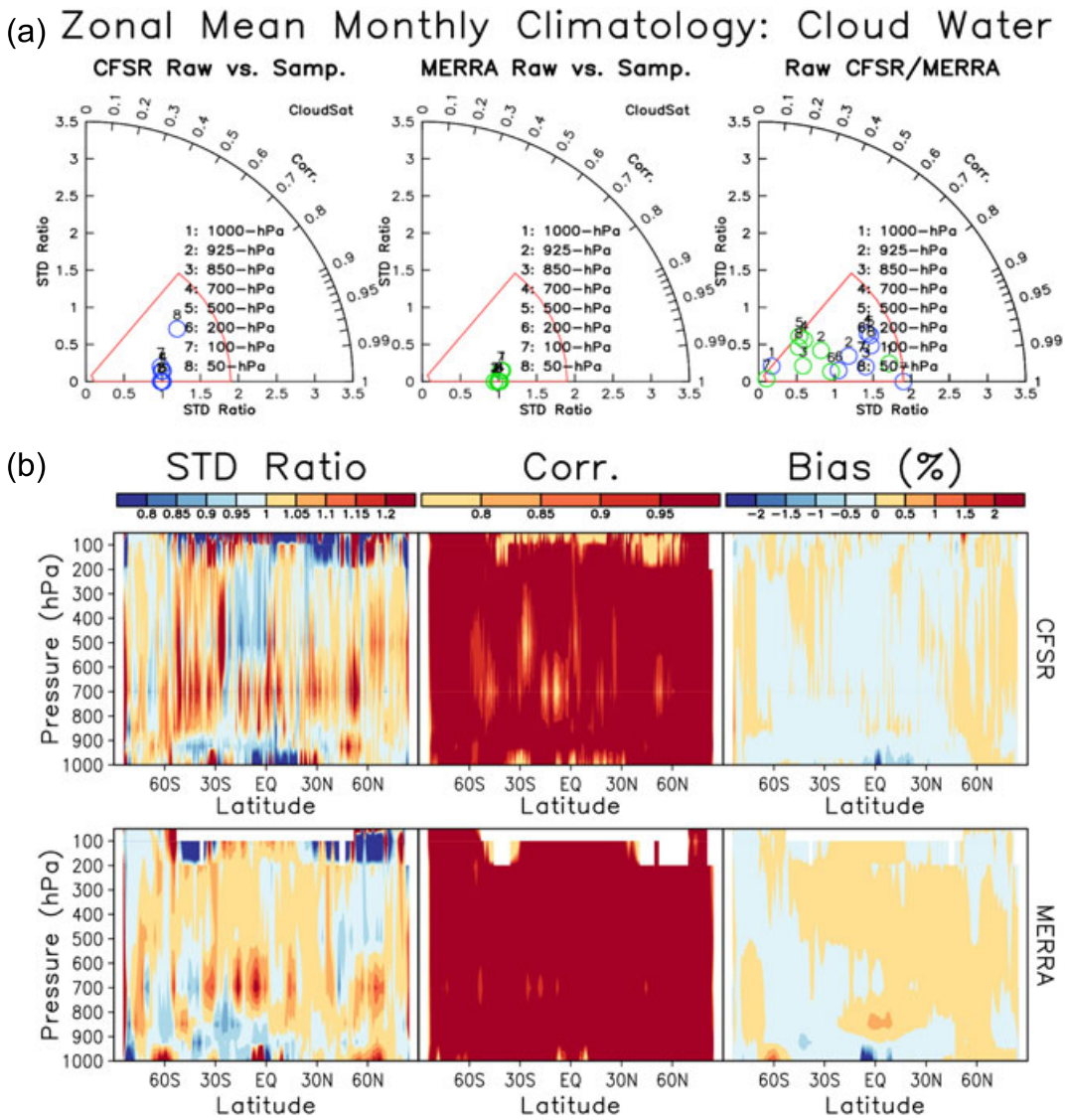

Figure 6. (a) As the 4th row in Figure 5 except for the zonal mean field. (b) Latitude-height cross sections of (left) standard deviation ratio, (center) correlation, and (right) normalized bias between zonal mean raw and satellite-sampled cloud water based on (upper) CFSR and (lower) MERRA.

upper levels. Latitudinally, the two differ with respect to the locations of the largest sampling effect.

3.3. Monthly Means

[18] Similar comparisons are done for monthly means of raw and sampled data (Figure 7). Compared to other variables, CFSR cloud water shows exceptionally large and significant sensitivity to orbital sampling, and the sensitivity is larger than in the case of monthly climatologies at all vertical levels (cf. Figure 5). Raw CFSR and MERRA also differ significantly, and by a larger extent than in the cases of long-term annual mean and monthly climatology. As in those cases, orbital sampling effect is consistently larger with CFSR.

[19] Again, zonal mean cloud water is examined. The overall sampling effect becomes much smaller and insignificant (Figure 8a). Latitude-height cross sections show that the orbital sampling effect is contributed by both standard deviation ratios (Figure 8b, left panels) and correlations (Figure 8b, center panels); more so for CFSR. Relatively large sensitivity to orbital sampling is seen for CFSR at almost all levels, while at selected levels for MERRA. The two also differ with respect to the latitudinal locations of the largest sampling effect.

3.4. Bias

[20] Bias of the sampled data normalized by the standard deviation of the long-term annual mean of the raw data is listed in Table 2. For the cases of monthly climatology and monthly means, percent biases are smaller since the denominators used for normalization are larger (not shown). Except for 2-hPa temperature and specific humidity, the bias of the sampled data is within ~3% of the standard deviation of the raw data. As such, the total RMS difference between raw and sampled data is largely determined by the degree of pattern similarity (as discussed in sections 3.1), not the difference in the overall magnitude. Relatively large percent biases (9–10%) for 2-hPa temperature and humidity is a

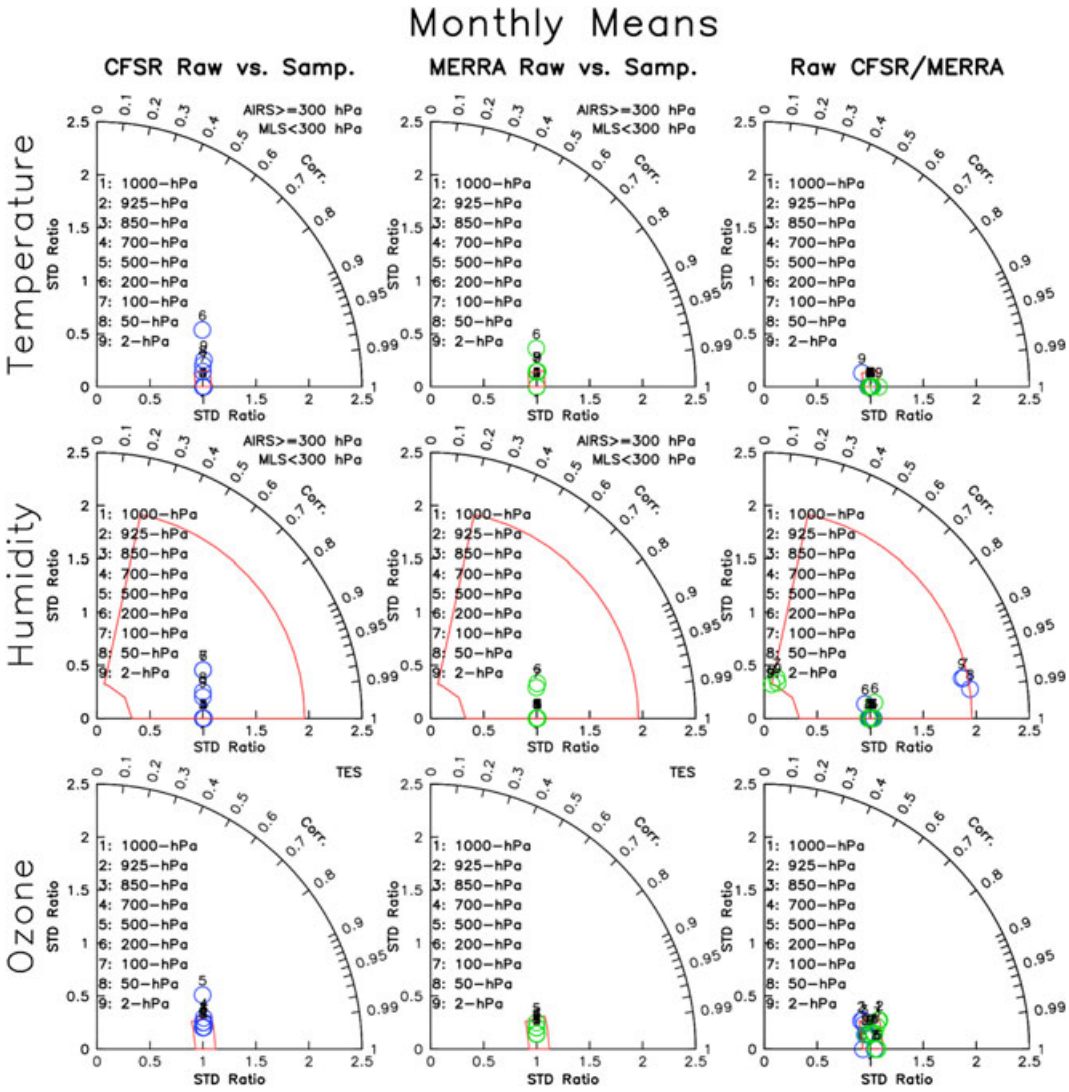

Figure 7. As Figure 3 except for monthly mean fields.

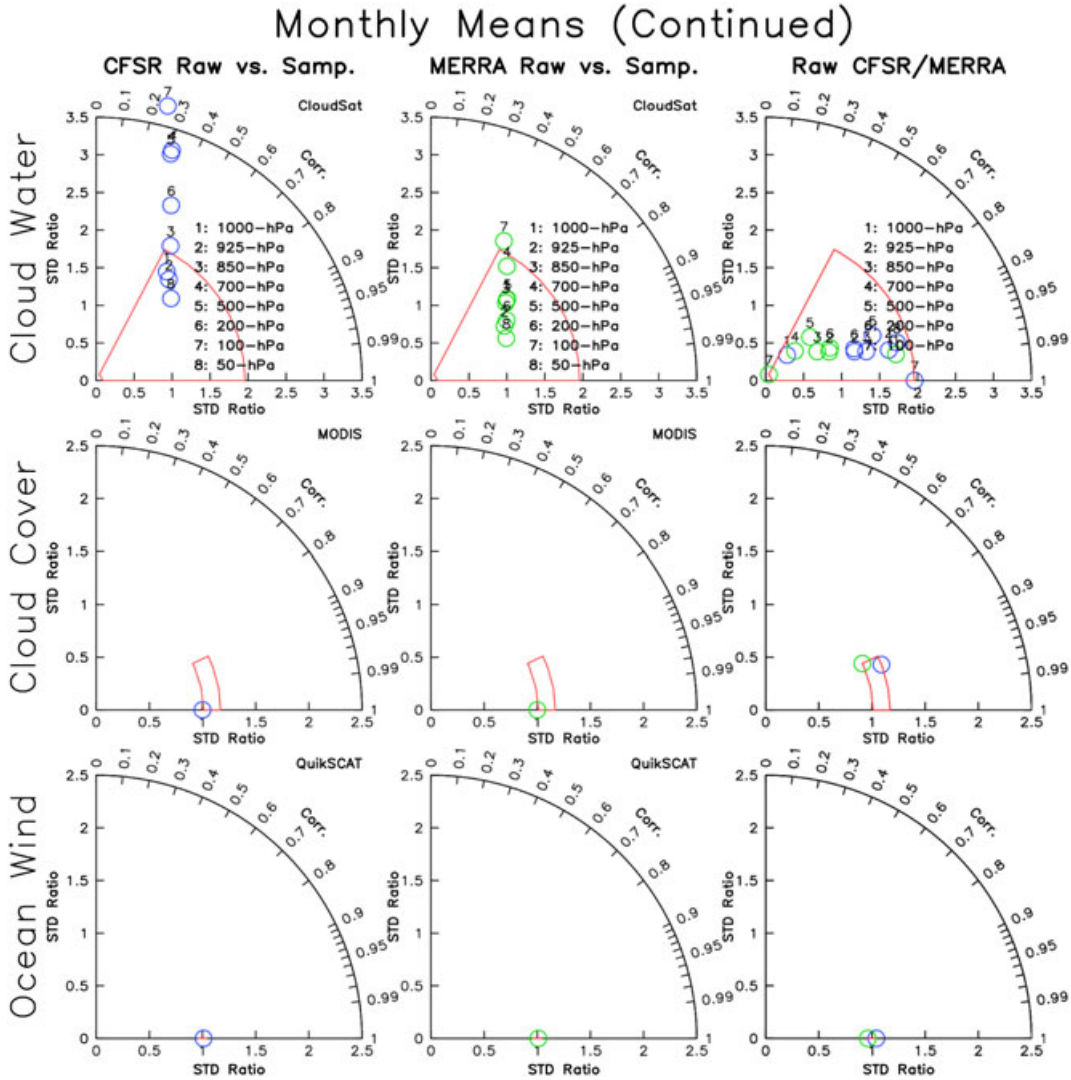

Figure 7. (continued)

result of relatively small standard deviation of the raw data at this level; in these cases, the absolute sampling biases are ~0.3 K and 1E−8 kg/kg, respectively, which are much smaller than the mean difference between raw data from the two reanalysis products themselves, i.e., the sampling biases are within the estimated raw data uncertainty. These results suggest that the bias of the sampled data is largely negligible. Similar conclusions have been made by *Aghedo et al*., [2011] for the TES instrument for zonal mean ozone, carbon monoxide, and water vapor based on four years of data from two climate chemistry models. Larger percent biases (~10–20%) were reported for the same instrument by *Luo et al*., [2002] for ozone and carbon monoxide from a climate chemistry model, on a single day basis. Using GCM, *Lin et al*., [2002] evaluated the orbital sampling effect for the TRMM satellite and found OLR sampling bias of less than 5%, but

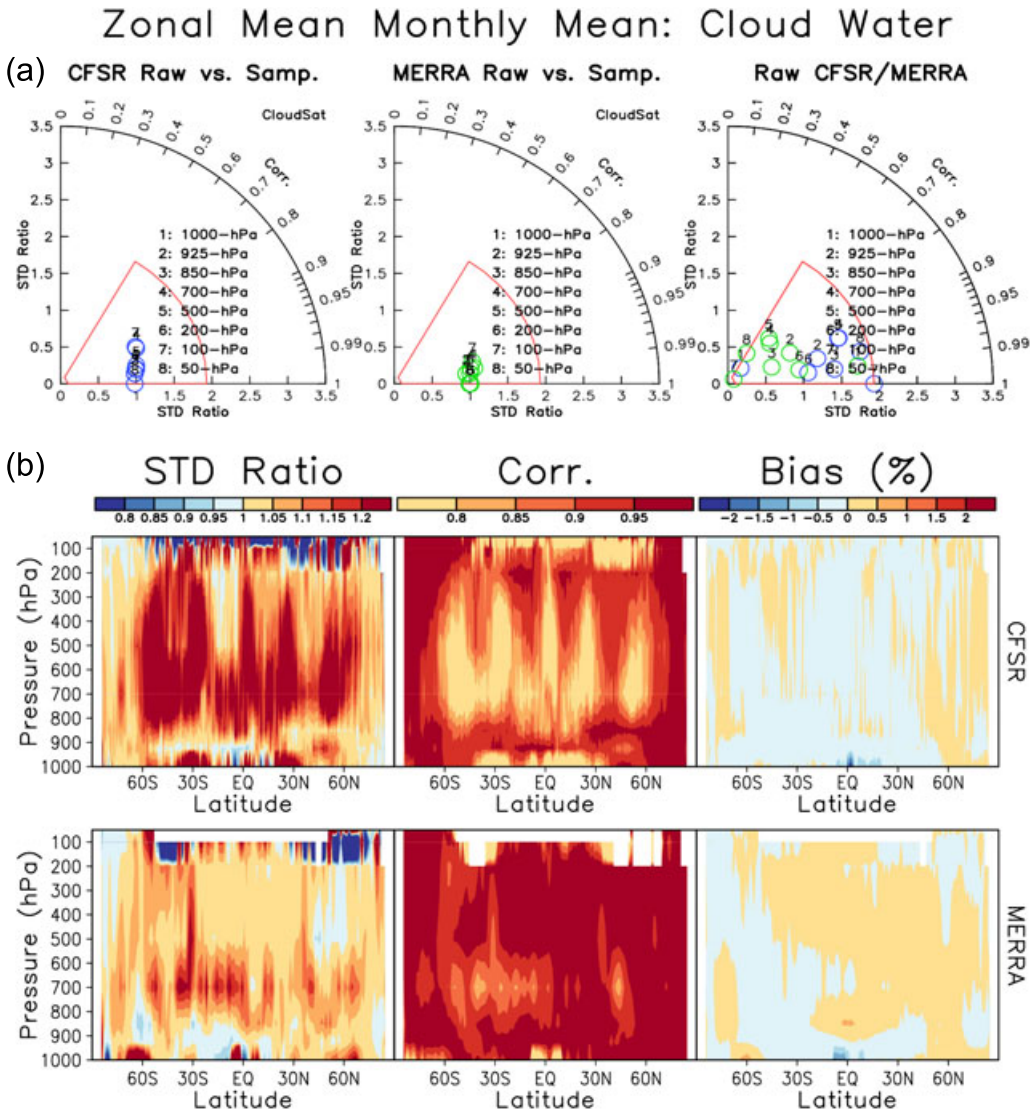

Figure 8. As Figure 6 except for monthly mean fields.

precipitation sampling bias of 20–40%, on a single month basis. *Engelen et al.*, [2000] found up to 5–7 K differences between GCM and satellite-sampled brightness temperatures for a single month. Relatively large biases were also shown by *Li et al.*, [2007] for MLS cloud ice based on one year of the European Centre for Medium-Range Weather Forecasts (ECMWF) analysis data. Comparison of these results suggests that the magnitude of the sampling bias is partly dependent on the time scale of interest.

[21] Spatial distribution of the sampling bias (not normalized) is shown in Figure 2 (right column). Land/sea dependence is most evident for 1000-hPa temperature, humidity, ozone, and column total cloud cover, due presumably to the different magnitudes of the diurnal cycle over land and sea. Latitudinal dependence is strongest for 200-hPa temperature and humidity, likely related to the gradients of these variables at different latitudes. Large-scale patterns characteristic of zonal wavenumber 4 are evident for ocean surface wind and column total cloud cover, which could be associated with atmospheric tides [*van den Dool et al.*, 1997].

4. Sensitivity to Spatiotemporal Resolution and Analysis Period

[22] In this section, we perturb a few aspects of the main analysis presented above and test the stability of the results obtained. For brevity, the tests focus on the case of monthly climatology, and on one variable, i.e., cloud water, which shows the largest orbital sampling effect relative to other variables and also the largest difference between the two reanalysis products. To be examined here is whether the sampling results are sensitive to the spatiotemporal resolution of the input data and the analysis period. The configuration for four test analyses are listed in Table 3 alongside two control analyses that have already been discussed in section 3.2.

[23] Results of the sensitivity tests are shown in Figure 9. For CFSR, regridding from the original $0.5° \times 0.5°$ grid to $1.25° \times 1.25°$ grid reduces the orbital sampling effect to largely within the observational uncertainty (Figure 9b), which also becomes nearly comparable to MERRA (Figure 9c). For MERRA, regridding from the original 1.25°

Table 2. Bias of the Sampled Data Normalized by the Standard Deviation of Long-Term Annual Mean of the Raw Data[a]

Variable	CFSR	MERRA
Temperature (≥300 hPa)	0.1–1.2%	0–0.5%
Temperature (<300 hPa)	−0.2–0.4%	−1.1–0.4%
	9.3% (0.28 K) at 2-hPa[b]	9.0% (0.34 K) at 2-hPa[b]
Specific Humidity (≥300 hPa)	−0.7–0.3%	−0.1–0.7%
Specific Humidity (<300 hPa)	−3.4–0.5%	−0.2–1.1%
	9.7% (1E−8 kg/kg) at 2-hPa[c]	
Ozone Mixing Ratio	−1.2–0.4%	−0.6–0%
Cloud Water Mixing Ratio	−2.5–1.3%	−3.1–2.0%
Column Total Cloud Cover	−2.2%	1.0%
Ocean Surface Wind	0.6%	−0.6%

[a]For variables with multiple vertical levels, bias is evaluated at each level, the range of which is indicated in the table.
[b]Mean difference between CFSR and MERRA raw temperature is 4.1 K at 2-hPa.
[c]Mean difference between CFSR and MERRA raw specific humidity is 3E−6kg/kg at 2-hPa.

1.25° grid to 8° Lon × 4° Lat grid further reduces the already insignificant sampling effect (Figure 9d), reassuring insignificant sampling effect in previous model evaluation studies based on the latter grid [e.g., *Li et al*., 2012a]. Reducing the temporal resolution from 3-hr to 6-hr barely changes the sampling effect (Figure 9e), and adding 10 years of data to the analysis only slightly changes the result (Figure 9f). These results suggest that the orbital sampling effect is not sensitive to reasonable changes in the temporal resolution of the input data or the length of the analysis period. In this regard, the relatively large sampling effect for CFSR compared to MERRA is attributed to the finer spatial resolution of CFSR. For instruments with small swaths, sampling frequency is sensitive to the spatial resolution of the field being sampled (Figure 1). At finer spatial resolutions, the percentage of the total number of reanalysis grid cells sampled by a given satellite instrument is reduced, leading to larger orbital sampling effect.

5. Summary and Discussion

[24] This paper evaluates the effect of orbital sampling on satellite–climate model comparisons. Six variables (temperature, specific humidity, ozone, cloud water, cloud cover, and ocean surface wind) associated with six satellite instruments (Aqua/AIRS, Aura/MLS, Aura/TES, CloudSat/CPR, Aqua/MODIS, and QuikSCAT/SeaWinds) are considered which represent the suite of satellite data being made available by NASA via the obs4MIPs project. Satellite sampling is done on 6-hourly CFSR and 3-hourly MERRA data, based on which monthly means, monthly climatologies, and long-term annual means are computed. Comparisons are then made between raw and sampled data for the three quantities, in terms of bias and pattern similarity. The results suggest that: (1) The bias introduced by orbital sampling is largely negligible, meaning that in nearly all cases, the errors introduced are within ~3% of the standard deviation of the raw data, and relatively large bias (9–10%) for 2-hPa temperature and specific humidity is still within the estimated observational uncertainty; (2) In terms of pattern similarity, cloud water and upper level specific humidity are the most sensitive to orbital sampling among the six variables considered, and the magnitude of the sampling effect is dependent on the spatial resolution—insignificant at 1.25° × 1.25° resolution for both variables; (4) For all variables considered, orbital sampling effects are not an important consideration for model evaluation at the resolution of 1.25° × 1.25°; (5) At the resolution of 0.5° × 0.5°, orbital sampling is potentially important for cloud water and upper level specific humidity when evaluating model long-term annual means and monthly climatologies, and for cloud water when evaluating monthly means, all in terms of pattern similarities; and not an important factor to consider for temperature, lower level specific humidity, ozone, column total cloud cover, and ocean surface wind; (6) Orbital sampling is not an important factor for evaluating zonal means in all cases considered. These results are summarized in Table 4. The two variables with the largest orbital sampling effect (cloud water and upper level specific humidity) are associated with two instruments (CloudSat/CPR and Aura/MLS) with small swaths (see Table 1), which, for variables with relatively large spatial heterogeneity, means significant undersampling over space, hence relatively large orbital sampling effect.

[25] Besides the orbital geometry considered here, satellite sampling is also affected by cloud filtering and other quality screenings [*Engelen et al*., 2000; *Aghedo et al*., 2011], which further reduces the sampling frequency. For variables based on visible light channels, the night hemisphere does not get sampled, which also reduces the sampling frequency. Reduced sampling can also be caused by mechanical,

Table 3. Configuration of Control and Test Analyses for Cloud Water Monthly Climatologies[a]

		Spatial Resolution	Temporal Resolution	Time Period
CFSR	Control	0.5° × 0.5°	6-hr	2000–2010
	Test	1.25° × 1.25°		
MERRA	Control	1.25° × 1.25°	3-hr	2000–2010
	Test 1	8° Lon × 4° Lat		
	Test 2		6-hr	
	Test 3			1990–2010

[a]Omitted for the test analyses are parameters identical to those used in the corresponding control analysis.

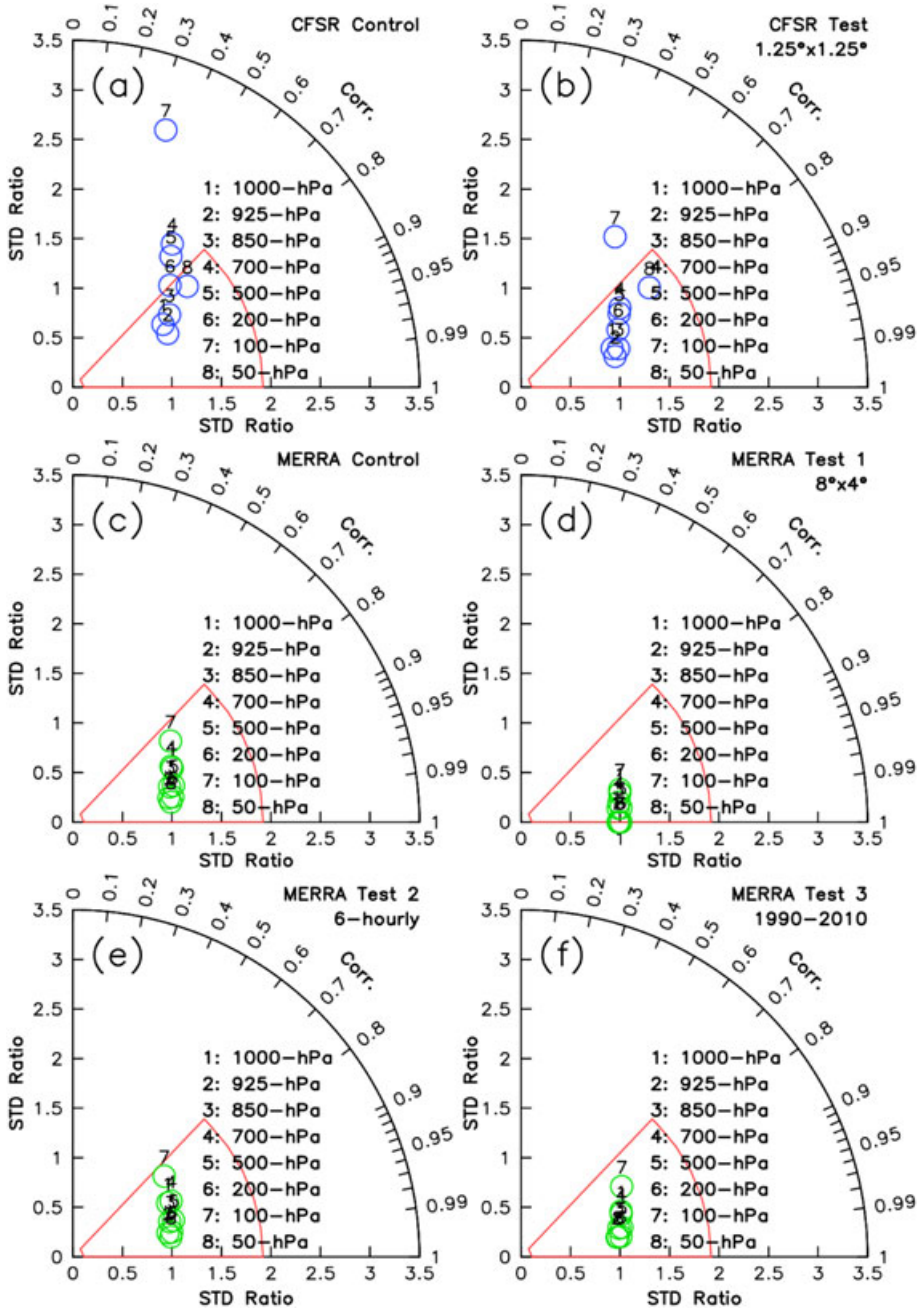

Figure 9. Taylor diagrams showing the results for the sensitivity tests listed in Table 3. The red "box" is as in the 4th row of Figure 5.

Table 4. The Need for Orbital Sampling or Not in Various Cases Based on the Results of the Current Study[a]

	0.5° × 0.5°			1.25° × 1.25°		
	Long-term Annual Mean	Monthly Clim.	Monthly Means	Long-term Annual Mean	Monthly Clim.	Monthly Means
Cloud Water	**Y**	**Y**	**Y**	N	N	N
Upper-level Humidity	**Y**	Y	N	N	N	N
Temperature, Lower-level Humidity, Ozone, Cloud Cover and Ocean Surface Wind	N	N	N	N	N	N
Zonal Mean of Any of the Above	N	N	N	N	N	N

[a]"**Y**" indicates orbital sampling is recommended, and "N" indicates orbital sampling is likely not necessary.

electronic, and other difficulties encountered by the satellite instrument. Other factors need to be considered when doing satellite–model comparisons include sensitivity and saturation level of the instrument [*Li et al*., 2007], errors of the retrieval algorithm and, for level 3 (i.e., gridded) product, errors introduced in mapping the swath data to gridded data [*Luo et al*., 2002]. For radiative fluxes, an important factor to consider is how "clear-sky" is defined. Instead of compositing over cloud-free scenes as used by satellite products, climate models in general produce clear-sky radiative fluxes by removing clouds and re-do the flux calculations under the same all-sky conditions, which can result in considerable differences compared to the satellite method [e.g., *Sohn et al*.,, 2010]. Regional scale analysis is needed to evaluate the sensitivity of the sampling effect to geographical locations (e.g., high vs. low latitudes, land vs. oceans) and seasons. Orbital sampling effect in the case of multiple satellites working in tandem [e.g., *Lee et al*., 2008] is worth investigation. Finally, a very important consideration is the size of the satellite sampling errors relative to the model–observation differences, the latter of which is often demonstrably larger, indicating less concern regarding the satellite sampling. Continued analyses are needed to better understand the uncertainties and limitations in satellite–model comparisons as observations and models improve.

[26] **Acknowledgments.** CFSR data were made available by NCEP and downloaded from the CISL Research Data Archive managed by NCAR. MERRA data were made available by the Global Modeling and Assimilation Office (GMAO) and disseminated by the GES DISC. DEW's, JLL's and BG's contribution to this study was carried out on behalf of the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

References

Aghedo, A. M., Bowman, K. W., Shindell, D. T., and G. Faluvegi, (2011), The impact of orbital sampling, monthly averaging and vertical resolution on climate chemistry model evaluation with satellite observations, *Atmos. Chem. Phys.*, *11*, 6493–6514, doi:10.5194/acp-11-6493-2011.

Bell, T. L., A. Abdullah, R. L. Martin, and G. R. North (1990), Sampling errors for satellite-derived tropical rainfall: Monte Carlo study using a space-time stochastic model, *J. Geophys. Res.*, *95*, 2195–2205, doi:10.1029/JD095iD03p02195.

Bell, T. L., and P. K. Kundu (1996), A study of the sampling error in satellite rainfall estimates using optimal averaging of data and a stochastic model, *J. Climate*, *9*, 1251–1268, doi:10.1175/1520-0442(1996)009<1251:ASOTSE>2.0.CO;2.

Engelen, R. J., L. D. Fowler, P. J. Gleckler, and M. F. Wehner (2000), Sampling strategies for the comparison of climate model calculated and satellite observed brightness temperatures, *J. Geophys. Res.*, *105*, 9393–9406, doi:10.1029/1999JD901182.

Gleckler, P., R. Ferraro, and D. Waliser (2011), Improving use of satellite data in evaluating climate models, *Eos Trans. AGU*, *92*, 172, doi:10.1029/2011EO200005.

Huffman, G. J., R. F. Adler, D. T. Bolvin, G. Gu, E. J. Nelkin, K. P. Bowman, Y. Hong, E. F. Stocker, and D. B. Wolff (2007), The TRMM Multi-satellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales, *J. Hydrometeor.*, *8*, 38–55, doi:10.1175/JHM560.1.

Jiang, J. H., et al. (2010), Five year (2004–2009) observations of upper tropospheric water vapor and cloud ice from MLS and comparisons with GEOS-5 analyses, *J. Geophys. Res.*, *115*, D15103, doi:10.1029/2009JD013256.

Jiang, J. H., et al. (2012), Evaluation of cloud and water vapor simulations in CMIP5 climate models using NASA "A-Train" satellite observations, *J. Geophys. Res.*, *117*, D14105, doi:10.1029/2011JD017237.

Laughlin, C. R. (1981), On the effect of temporal sampling on the observation of mean rainfall, in: Precipitation Measurements From Space, Workshop Report, edited by D. Atlas and O. W. Thiele, pp. D59–D66, NASA Goddard Space Flight Cent., Greenbelt, Md.

L'Ecuyer, T. S., and J. Jiang, (2010), Touring the atmosphere aboard the A-Train, *Phys. Today*, *63*, 36–41.

Lee, T., O. Wang, W. Tang, and W. T. Liu (2008), Wind stress measurements from the QuikSCAT-SeaWinds scatterometer tandem mission and the impact on an ocean model, *J. Geophys. Res.*, *113*, C12019, doi:10.1029/2008JC004855.

Li, J.-L., et al. (2005), Comparisons of EOS MLS cloud ice measurements with ECMWF analyses and GCM simulations: Initial results, *Geophys. Res. Lett.*, *32*, L18710, doi:10.1029/2005GL023788.

Li, J.-L. F., et al. (2012a), An observationally-based evaluation of cloud ice water in CMIP3 and CMIP5 GCMs and contemporary reanalyses using contemporary satellite data, *J. Geophys. Res.*, doi:10.1029/2012JD017640, in press.

Li, J.-L. F., S. Lee, D. E. Waliser, S. Lee., B. Guan, G. Stephens, J. Teixeira (2012b), An observation-based evaluation of cloud liquid water in CMIP3 and CMIP5 GCMs and contemporary analyses, *J. Geophys. Res.*, in revision.

Li, J.-L., J. H. Jiang, D. E. Waliser, and A. M. Tompkins (2007), Assessing consistency between EOS MLS and ECMWF analyzed and forecast estimates of cloud ice, *Geophys. Res. Lett.*, *34*, L08701, doi:10.1029/2006GL029022.

Li, Q., R. L. Bras, and D. Veneziano (1996), Analysis of Darwin rainfall data: Implications on sampling strategy, *J. Appl. Meteorol.*, 35372–35385.

Lin, X., L. D. Fowler, and D. A. Randall (2002), Flying the TRMM Satellite in a general circulation model, *J. Geophys. Res.*, *107*, D16, doi:10.1029/2001JD000619.

Luo, M., R. Beer, D. J. Jacob, J. A. Logan, and C. D. Rodgers (2002), Simulated observation of tropospheric ozone and CO with the Tropospheric Emission Spectrometer (TES) satellite instrument, *J. Geophys. Res.*, *107*, D15, doi:10.1029/2001JD000804.

North, G. R. (1988), Survey of sampling problems for TRMM, in: Tropical Rainfall Measurements. J. S. Theon and N. Fugono (Eds.), Deepak Publishing, Hampton, VA. 337–348.

North, G. R., S. S. P. Shen, and R. Upson (1993), Sampling errors in rainfall estimates by multiple satellites, *J. Appl. Meteor.*, *32*, 399–410.

Randall, D. A., et al. (2007), Cilmate models and their evaluation, in: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor and H. L. Miller (eds.)], Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Rienecker, M. M., et al. (2011), MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications, *J. Climate*, *24*, 3624–3648, doi:10.1175/JCLI-D-11-00015.1.

Saha, S., et al. (2010), The NCEP Climate Forecast System Reanalysis, *Bull. Amer. Meteor. Soc.*, *91*, 1015–1057, doi:10.1175/2010BAMS3001.1.

Sohn, B. J., T. Nakajima, M. Satoh, and H.-S. Jang (2010), Impact of different definitions of clear-sky flux on the determination of longwave cloud radiative forcing: NICAM simulation results, *Atmos. Chem. Phys.*, *10*, 11641–11646, doi:10.5194/acp-10-11641-2010.

Soman, V. V., J. B. Valdés, and G. R. North (1996), Estimation of sampling errors and scale parameters using two- and three-dimensional rainfall data analyses, *J. Geophys. Res.*, *101*, 26453–26460, doi:10.1029/96JD01387.

Su, H., D. E. Waliser, J. H. Jiang, J. Li, W. G. Read, J. W. Waters, and A. M. Tompkins (2006), Relationships of upper tropospheric water vapor, clouds and SST: MLS observations, ECMWF analyses and GCM simulations, *Geophys. Res. Lett.*, *33*, L22802, doi:10.1029/2006GL027582.

Taylor, K. E. (2001), Summarizing multiple aspects of model performance in a single diagram, *J. Geophys. Res.*, *106*, 7183–7192, doi:10.1029/2000JD900719.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl (2012), An overview of CMIP5 and the experiment design, *Bull. Amer. Meteor. Soc.*, *93*, 485–498. doi:10.1175/BAMS-D-11-00094.1.

Teixeira, J., D. Waliser, R. Ferraro, P. Gleckler, and G. Potter (2011), Satellite observations for CMIP5 simulations, CLIVAR Exchanges No. 56, Vol. 16, No.2, May 2011.

van den Dool, H. M., S. Saha, J. Schemm, and J. Huang (1997), A temporal interpolation method to obtain hourly atmospheric surface pressure tides in reanalysis 1979–1995, *J. Geophys. Res.*, *102*, 22013–22024.

Waliser, D. E., et al. (2009), Cloud ice: A climate model challenge with signs and expectations of progress, *J. Geophys. Res.*, *114*, D00A21, doi:10.1029/2008JD010015.

Waliser, D. E., J.-L. F. Li, T. S. L'Ecuyer, and W.-T. Chen (2011), The impact of precipitating ice and snow on the radiation balance in global climate models, *Geophys. Res. Lett.*, *38*, L06802, doi:10.1029/2010GL046478.

Williams, D. N., et al. (2009), The Earth System Grid: Enabling Access to Multimodel Climate Simulation Data, *Bull. Amer. Meteor. Soc.*, *90*, 195–205, doi: http://dx.doi.org/10.1175/2008BAMS2459.1.